ultimately will lead to the development of a new generation of drugs. The availability of gene sequence information is changing the way biomedical research is conducted and the rate of discovery. Having the sequence of a genome, however, does not reveal what the genes do nor how the encoded proteins function, how cells and tissues develop, nor give insights on the etiology and cure of diseases. Before the fruits of the information obtained by sequencing a genome can be realized, encoded proteins and their functions must be identified.

Hence, there is an emergence of proteomics in which the challenge is to unravel the plethora of information that has been obtained by virtue of sequencing of the human genome and other genomes. The focus is assigning functions to genes that have been identified by sequence. It is, however, a simpler task to identify a gene by sequencing it than it is to discover a function of the gene or the encoded protein. Various approaches, including biochemical, genetic and informatics approaches, to identifying proteins encoded by genes have been pursued in the attempt to do this. Informatics approaches attempt to define gene functions based on computer searches that compare gene sequences with the sequences of genes that encode proteins with known or purportedly known functions. Because of the discontinuity between gene sequence and function, these approaches have had limited success. Defining gene functions remains dependent on traditional approaches of genetics and biochemistry. The genetic approach is based on disrupting a genes function and then observing the effects of that disruption; the biochemical approach is based on correlating biochemical changes with function. To make any headway, high throughput analyses are required.

Please replace the paragraph on page 4, lines 13-30, with the following paragraph:

SUMMARY OF THE INVENTION

Provided herein are methods and products for screening and identifying molecules, particularly proteins and nucleic acids, from among large collections. In particular, collections of capture agents (*i.e.*, receptors, such as antibodies or other receptors) that specifically bind to identifiable protein binding partners, designated polypeptide tags herein, in which each capture agent has been selected or designed to bind with high selectivity and specificity to a pre-selected polypeptide tag, such as an epitope or ligand or portion thereof are provided. The collections, which contain identifiable capture agents, such as antibodies, are provided in any suitable format, including liquid phase and solid phase formats, as long as the capture agents, such as antibodies are identifiable (addressable). Addressable arrays of the capture agents are exemplified herein. The methods herein exemplified with respect to arrays can be practiced with any other format, including capture agents, such as antibodies, linked to RF tags, detectable beads, bar coated beads and other such formats. The collections serve as devices to sort, and ultimately, identify, proteins and genes and other molecules of interest.

Please replace the paragraph on page 5, lines 21-27, with the following paragraph:

The addressable capture agent, such as an antibody, collections provide a means to sort molecules tagged with the sequence of amino acids of the polypeptide that specifically reacts with the capture agent. The sorting relies on the highly specific interaction between capture agents, such as antibodies, in the collection and the polypeptide tags, such as epitope tags, that are introduced into collections of molecules to be sorted.

Please replace the paragraph on page 6, lines 7-16, with the following paragraph:

In practice proteins tagged with the polypeptide tags are bathed over an array of capture agents or reacted with the collection of capture agents linked to

identifiable supports, such as beads, under suitable binding conditions. By virtue of the binding specificity of the preselected tags for particular capture agents, the proteins are sorted according their preselected tag. The identity of the tag is then known, since it reacts with a particular capture agent whose identity is known by virtue of its position in the array or its identifier, such as its linkage to an optically coded, including as color coded or bar coded, or an electronically-tagged, such as a microwave or radio frequency (RF)-tagged, particle.

A4 cont.

Please replace the paragraph beginning on page 6, line 30, through page 7, line 5, with the following paragraph:

In another embodiment, the solid supports constitute coded particles (beads), such as microspheres that can be handled in liquid phase and then layered into a two dimensional array. The particles, such as microspheres, are encoded optically, such as by color or bar coded, chemically coded, electronically coded or coded using any suitable code that permits identification of the bead and capture agent bound thereto. The capture agent is coated on or otherwise linked to the support.

A5

Please replace the paragraphs beginning on page 8, line 13, through page 9, line 24, with the following paragraphs:

Methods using the capture agent (also referred to herein as a receptor) collections, such as antibody collections, for sorting molecules labeled with the binding pair, such as an epitope, tags are provided. The methods include the steps of creating a master tagged library by adding nucleic acids encoding the tags; dividing a portion of the master library into N reactions; amplifying each reaction with the nucleic acid encoding the divider sequences and translating to produce N translated reactions mixtures; reacting each of the reactions mixtures with one collection of the antibodies, using for example conditions used for western blotting; identifying the proteins of interest by a suitable screen,

A6

thereby identifying the particular polypeptide tag on the protein by virtue of the capture agent which the protein of interest binds.

The first sort is designed to reduce diversity by a significant factor. Standard screening methods may then be employed to screen the new sublibrary. If a further reduction in diversity is desired a second sort can be performed. By appropriate selection of the number of antibodies (or other receptors), the number of D's and pools and the number of collections in the first screen, the optional second screen can be designed so that the resulting collection should contain only a single protein or only a small number of proteins.

A16 cont

A second sort starting from the nucleic acid reaction mixture that contains the nucleic acid from which the protein of interest was translated can be performed. In this step, a new set of the polypeptide tags is added to the nucleic acid by amplification or ligation followed by amplification. Prior to or simultaneously with this, the nucleic acid encoding the prior polypeptide tag, such as epitope tag, is removed either by cleavage, such as with a restriction enzyme or by amplification with a primer that destroys part or all of the epitope-encoding nucleic acid. The new tags are added, and resulting nucleic acids are translated and are reacted with a single addressable collection of antibodies. The proteins sort according to their polypeptide tag, and a screen is run to identify the protein of interest. At this point, the diversity of the molecules at the addressable locus of the antibody collection should be 1 (or on the order of 1 to 10). The nucleic acids that contain the protein of interest are then amplified with a tag that amplifies nucleic acid molecules that contain nucleic acids encoding the identified polypeptide tag, to thereby produce nucleic acid encoding a protein of interest. The primer for amplification, particularly in methods in which a second or additional sorting steps are contemplated, can include all or only a sufficient portion of the tag to serve as a primer to thereby remove at least part of the "E" portion of the polypeptide tag from the encoded protein.

Pl as replace th paragraphs beginning n pag 10, line 4, through pag 12, line 14, with the following paragraphs:

In exemplified embodiments, the master libraries are complementary DNA (cDNA) libraries, and the polypeptide tags are encoded by primers or oligonucleotides that are introduced into the cDNA molecules in the library. In the first step in these methods, a master collection of nucleic acids, which each include, generally at one end, such as at the 3'-end or 5'- end of the nucleic acid molecule, nucleic acid encoding a preselected polypeptide containing an epitope (*i.e.*, specific sequence of amino acids required for specific binding to the capture agent), is prepared. Samples from the master collection are divided into N pools, such as 50, 100, 200, 250 (or conveniently 96 or a multiple (96, 96 x 1, 96 x 2 . . . n, wherein n is 1 to as many pools as needed, such as 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 300, 500, 10^r, where r is 2 or more, thereof). In each pool one of the n divider sequences (D_n) is used to amplify all nucleic acids that include that particular D.

Each amplified pool is translated and the proteins contained therein are contacted with one of the capture agent collections, such as antibody collections, in which the tag for which each capture agent is specific and is known, such as by virtue of its position in an addressable two or three-dimensional array or its linkage to an identifiable particulate support. After contacting, capture agent-protein complexes are identified using standard methods, such as an assay specific for the protein(s) of interest, or by addition of other suitable reagents. Colorimetric, luminescent, fluorescent and other such assays are among the screening assays contemplated. By identifying the capture agent, *i.e.*, antibody, to which the protein of interest binds and the pool containing such capture agent, the original D_n pool is known as well as the epitope in the pool and diversity is reduced by n x m. A set of primers containing a portion of the epitope, designated FA, and including all of the E's, is used to amplify the D_m pool. This specifically amplifies only members of the pool that include the identified E tag, destroys the epitope in the translated

protein and introduces a new set of polypeptide tags encoding nucleic acid molecules into the pool, which is then translated and contacted with a single collection of antibodies; the collection is screened to identify complexes. Amplification of the nucleic acid encoding the identified E tag with a primer containing FB, where FB is all or a portion of the epitope, followed by translation results in a sample containing the protein(s) of interest.

If further reduction in diversity is desired, additional sorting steps may be employed using M_i and N_i tags, where "i" refers to the sorting step number and signifies that M and N may be different at each step. Each M and N can be selected to achieve the desired reduction in diversity. The diversity of the library = Div, is the number of different genes or proteins in a library; N_i is the number of divider sequences (each divider sequence is designated D_n) used in a particular sorting step, wherein n is from 2 up to N, typically at least about 10 to N_i x M_i, is the number of polypeptide tags, M_i is the number of different capture agents, such as antibodies and/or other receptors or portions thereof, in a collection, and each polypeptide tag is designated E_m, where m is 2 to M_i, preferably at least about 10 to M, and i is from 1 to Q, and Q is the number of sorting steps with the antibody collection. In particular, the diversity of the library (Div), Div = $(N_i \times Mi)(N_{i+1} \times M_{i+1}) \ldots (N_Q \times M_Q)$ where i, the sorting step is 1 to Q. If N, $N_i \ldots N_Q$ are the same number at each step, and M, $M_i \ldots M_Q$ are the same number at each step, the DIV = $(N \times M)^Q$. If the goal is to reduce diversity to a desired level, such as 1, then $Div/(N_i \times M_i)(N_{i-1} \times M_{i-1}) \ldots (N_Q \times M_Q)$ = the desired level of diversity, and M and N at each sort should be selected accordingly.

A7 cont.

Hence, for example, if there are 10^6 proteins in a library, if there are 100 different antibodies in each collection (M), and 100 replicate antibody collections are used (N), and there are two (Q = 2) sorting steps, then for a library with a diversity of 10^6 (Div), the number of reactions into which the initial master collection is divided, will be 100. Generally the number of sorts is one or two. It can be more, but the last step is designed so that at this step substantially all

of the molecules at a locus are the same. Alternatively, there may be fewer sorting steps, typically one, which substantially reduce the diversity. Other screening methods can be used in place of further sorting steps to identify proteins corresponding to library members of interest. In this example, after the first sort, the diversity is reduced such that a protein corresponding to library member of interest is present at about 1 in 100; diversity (DIV) has been reduced by a factor of 10^4. Rather than perform a second sort, other screening methodologies can be used to identify the desired one amongst 100.

Please replace the paragraph on page 13, lines 17-27, with the following paragraph:

In one embodiment, combinations of the collections of capture agents, such as antibodies and oligonucleotides that encode polypeptide epitopes to which the capture agents selectively bind are provided. Kits containing the oligonucleotides and capture agents, such as antibodies, and optionally containing instructions and/or additional reagents are provided. The combinations include a collection of capture agents, antibodies, that specifically bind to a set of preselected epitopes, and a set of oligonucleotides that encode each of the epitopes. The oligonucleotides are single-stranded, double-stranded or include double-stranded and single-stranded portions, such as single-stranded overhangs created by restriction endonuclease cleavage.

Please replace the paragraph on page 14, lines 3-5, with the following paragraph:

FIGURE 2 also illustrates nested sorting; this sort is identical to the sort illustrated in Fig 1 except that the F2 and F3 sublibraries have been arranged into arrays.

Please replace th paragraph on page 14, lin 14, with the following paragraph:

FIGURE 7 depicts an alternative scheme using linker addition.

Q10

~~Please replace~~ the paragraph on page 15, lines 16-25, with the following paragraph:

DETAILED DESCRIPTION

A. DEFINITIONS

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as is commonly understood by one of skill in the art to which this invention belongs. In the event there are different definitions for terms herein, the definitions in this section control. Where permitted, all patents, applications, published applications and other publications and sequences from GenBank and other databases referred to throughout in the disclosure herein are incorporated by reference in their entirety.

Q11

Please replace the paragraphs beginning on page 16, line 15, through page 18, line 28, ~~with the following~~ paragraphs:

As used herein, an epitope tag refers to a sequence of amino acids that includes the sequence of amino acids, herein referred to as epitope, to which an anti-tag capture agent, such as an antibody specifically binds. For polypeptide and epitope tags, the specific sequence of amino acids to which each binds is referred to herein generically as an epitope. Any sequence of amino acids that binds to a receptor therefor is contemplated. For purposes herein the sequence of amino acids of the tag, such as epitope portion of the epitope tag, that specifically binds to the capture agent is designated "E", and each unique epitope is an E_m. Depending upon the context "E_m" can also refer to the sequences of nucleic acids encoding the amino acids constituting the epitope. The polypeptide tag, such as epitope tag, may also include amino acids that are encoded by the divider region. In particular, the epitope tag is encoded by the oligonucleotides provided herein, which are used to introduce the tag. When reference is made to an epitope tag (i.e. binding pair for a particular receptor or

Q12

portion thereof) with respect to a nucleic acid, it is nucleic acid encoding the tag to which reference is made. For simplicity each polypeptide tag is referred to as E_m; when nucleic acids are being described the E_m is nucleic acid and refers to the sequence of nucleic acids that encode the epitope; when the translated proteins are described, E_m refers to amino acids (the actual epitope). The number of E's corresponds to the number of antibodies in an addressable collection. "m" is typically at least 10, more preferably 30 or more, more preferably 50 or 100 or more, and can be as high as desired and as is practical. Most preferably "m" is about a 1000 or more.

As used herein, D_n refers to each divider sequence. As described herein in certain embodiments in which division is effected by other methods D_n is optional. As with each E_m the D_n is either nucleic acid or amino acids depending upon the context. Each D_n is a divider sequence that is encoded by a nucleic acid that serves as a priming site to amplify a subset of nucleic acids. The resulting amplified subset of nucleic acids contains all of the collection of E_m sequences and the D_n sequences used as a priming site for the amplification. As described herein, the nucleic acids include a portion, preferably at the end, that encodes each E_mD_n. Generally the encoding nucleic acid is 5'- E_m-D_n -3' on the nucleic acid molecules in the library. D is an optional unique sequence of nucleotides for specific amplification to create the sublibraries. For large libraries, the original library can be divided into sublibraries and then the tag-encoding sequences added, rather than adding the tag-encoding sequences to the master library. The size of D is a function of the library to be sorted, since the larger the library the longer the sequence needed to specify a unique sequence in the library. Generally D, depending upon the application, should be at least 14 to 16 nucleic acid bases long and it may or may not encode a sequence of amino acids, since its function in the method is to serve as a priming site for PCR amplification; D is 2 to n, where n is 0 or is any desired number and is generally 10 to 10,000, 10 to 1000, 50 to 500, and about 100 to 250. The number of D can be as high as 10^6 or higher. The divider

sequences D are used to amplify each of the "n" samples from the tagged master library, and generally is equal to the number of antibody collections, such as arrays, used in the initial sort. The more collections (divisions) in the initial screen, the lower diversity per addressable locus. The initial division number is selected based upon the diversity of the library and the number of capture agents. The more E's, the fewer D's are needed, and vice versa, for a library having a particular diversity (Div). As used herein, diversity (Div) refers to the number of different molecules in a library, such as a nucleic acid library. Diversity is distinct from the total number of molecules in any library, which is greater. The greater the diversity, the lower the number of actual duplicates there are. Ideally the (number of different molecules)/(total molecules) is approximately 1. If the number of molecules that are randomly tagged to create the master library is less than the initial diversity, then statistically each of the molecules in the master library should be different.

A12 cont.

As used herein, an array refers to a collection of elements, such as antibodies, containing three or more members. An addressable array is one in which the members of the array are identifiable, typically by position on a solid phase support or by virtue of an identifiable or detectable label, such as by color, fluorescence, electronic signal (*i.e.* RF, microwave or other frequency that does not substantially alter the interaction of the molecules of interest), bar code or other symbology, chemical or other such label. Hence, in general the members of the array are immobilized to discrete identifiable loci on the surface of a solid phase or directly or indirectly linked to or otherwise associated with the identifiable label, such as affixed to a microsphere or other particulate support (herein referred to as beads) and suspended in solution or spread out on a surface.

Please replace the paragraphs beginning on page 19, line 23, through page 20, line 17, with the f llowing paragraphs:

As used herein, matrix or support particles refers to matrix materials that are in the form of discrete particles. The particles have any shape and dimensions, but typically have at least one dimension that is 100 mm or less, 50 mm or less, 10 mm or less, 1 mm or less, 100 μm or less, 50 μm or less and typically have a size that is 100 mm^3 or less, 50 mm^3 or less, 10 mm^3 or less, and 1 mm^3 or less, 100 μm^3 or less and may be on order of cubic microns. Such particles are collectively called "beads."

As used herein, a capture agent, which is used interchangeably with a receptor, refers to a molecule that has an affinity for a given ligand or a with a defined sequence of amino acids. Capture agents may be naturally-occurring or synthetic molecules, and include any molecule, including nucleic acids, small organics, proteins and complexes that specifically bind to specific sequences of amino acids. Capture agents are receptors may also be referred to in the art as anti-ligands. As used herein, the terms, capture agent, receptor and anti-ligand are interchangeable. Capture agents can be used in their unaltered state or as aggregates with other species. They may be attached or in physical contact with, covalently or noncovalently, a binding member, either directly or indirectly via a specific binding substance or linker. Examples of capture agents, include, but are not limited to: antibodies, cell membrane receptors surface receptors and internalizing receptors, monoclonal antibodies and antisera reactive or isolated components thereof with specific antigenic determinants (such as on viruses, cells, or other materials), drugs, polynucleotides, nucleic acids, peptides, cofactors, lectins, sugars, polysaccharides, cells, cellular membranes, and organelles.

Please replace the paragraphs beginning on page 20, line 30, through page 22, line 2, with the following paragraphs:

As noted, contemplated herein, are pairs of molecules, generally proteins that specifically bind to each other. One member of the pair is a polypeptide that is used as a tag and encoded by nucleic acids linked to the library; the

other member is anything that specifically binds thereto. The collections of capture agents, include receptors, such as antibodies or enzymes or portions thereof and mixtures thereof that specifically bind to a known or knowable defined sequence of amino acids that is typically at least about 3 to 10 amino acids in length.

As used herein, antibody refers to an immuoglobulin, whether natural or partially or wholly synthetically produced, including any derivative thereof that retains the specific binding ability of the antibody. Hence antibody includes any protein having a binding domain that is homologous or substantially homologous to an immunoglobulin binding domain. For purposes herein, antibody includes antibody fragments, such as Fab fragments, which are composed of a light chain and the variable region of a heavy chain. Antibodies include members of any immunoglobulin class, including IgG, IgM, IgA, IgD and IgE. Also contemplated herein are receptors that specifically binding to a sequence of amino acids.

Hence for purposes herein, any set of pairs of binding members, referred to generically herein as a capture agent/polypeptide tag, can be used instead of antibodies and epitopes per se. The methods herein rely on the capture agent/polypeptide tag, such as an antibody/epitope tag, for their specific interactions, any such combination of receptors/ligands (epitope tag) can be used. Furthermore, for purposes herein, the capture agents, such as antibodies employed, can be binding portions thereof.

As used herein, antibody fragment refers to any derivative of an antibody that is less than full length, retaining at least a portion of the full-length antibody's specific binding ability. Examples of antibody fragments include, but are not limited to, Fab, Fab', $F(ab)_2$, single-chain Fvs (scFv), Fv, dsFv diabody and Fd fragments. The fragment can include multiple chains linked together, such as by disulfide bridges. An antibody fragment generally contains at least about 50 amino acids and typically at least 200 amino acids.

Please replace the paragraph on page 22, lines 6-7, with the following paragraph:

As used herein, a dsFv refers to an Fv with an engineered intermolecular disulfide bond, which stabilizes the V_H-V_L pair.

Q15

Please replace the paragraph beginning on page 24, line 30, through page 25, line 2, with the following paragraph:

As used herein, oligonucleotides refer to polymers that include DNA, RNA, nucleic acid analogs, such as PNA, and combinations thereof. For purposes herein, primers and probes are single-stranded oligonucleotides.

Q16

Please replace the paragraph on page 27, lines 5-11, with the following paragraph:

As used herein, suitable conservative substitutions of amino acids are known to those of skill in this art and may be made generally without altering the biological activity of the resulting molecule. Those of skill in this art recognize that, in general, single amino acid substitutions in non-essential regions of a polypeptide do not substantially alter biological activity (see, *e.g.*, Watson *et al. Molecular Biology of the Gene*, 4th Edition, 1987, The Benjamin/Cummings Pub. co., p.224).

Q17

Please replace the paragraph on page 28, lines 8-18, with the following paragraph:

The methods and collections herein are described and exemplified with particular reference to antibody capture agents, and polypeptide tags that include epitopes to which the antibodies bind, but is it to be understood that the methods herein can be practiced with any capture agent and any polypeptide tag therefor. It also to be understood that combinations of collections of any

Q18

capture agents and polypeptide tag therefor are contemplated for use in any of the embodiments described herein. It is also to be understood that reference to array is intended to encompass any addressable collection, whether it is in the form of a physical array or labeled collection, such as capture agents bound to colored beads.

Q18 Com.

Please replace the paragraph on page 29, lines 3-20, with the following paragraph:

1. Primers

Provided herein are sets of nucleic acid molecules that are primers or double-stranded oligonucleotides, which are double-stranded versions of the primers, and combinations of sets of primers and/or double-stranded oligonucleotides. The selection of single-stranded or double-stranded primers the use in the various steps of the methods provided herein and/or depends upon the embodiment employed. The primers, which are employed in some of the embodiments of the methods for tagging molecules, are central to the practice of such methods. The primers contain oligonucleotides, which include the formulae as depicted in Figure 9. The primers and double-stranded oligonucleotides may include restriction site(s) and for targeted amplifications, as exemplified below for example for antibody libraries, of sufficient portions of genes of interest. These primers may be forward or reverse primers, where the forward primer is that used for the first round in a PCR amplification.
The primers, described below and depicted in the figure, are provided as sets. Also provided are combinations of one or more of each set. The primers are central to the methods provided herein.

Q19

Please replace the paragraph beginning on page 29, line 27, through page 30, line 12, with the following paragraph:

Fig 9 illustrates the physical elements for construction of a tagged library and use of the addressable anti-tag antibody collections for identification of

Q20

genes (proteins) of interest. Four oligonucleotide/primer sets are provided in addition to the addressable collections, which for exemplification purposes are provided as arrays, an imaging system or reader to analyze the arrays and, optionally software to manage the information collected by the reader. In the embodiment depicted, the primer sets include E_mD_nC, where C is a portion in common amongst all of the oligonucleotides and can serve as a region for amplification of all tagged nucleic acids with differing E and/or D sequences (*e.g.*, D_1 thru D_n; E_1 thru E_m); DC, with differing D sequences (D_1 thru D_n), and an optional C, for common region, FAEC, with differing FA sequences (*e.g.*, FA_1 thru FA_n); and FBC, with differing FB sequences (*e.g.*, FB_1 thru FB_n). Each FA includes a portion of each epitope and can serve as a primer to amplify nucleic acids that encode a corresponding E_m, but the resulting amplified nucleic acids does not include the E_m epitope. FB_n is similar to FA_n, except that it can include E_n, if it is desired to retain the epitope.

Please replace the paragraph beginning on page 32, line 17, through page 33, line 2, with the following paragraph:

1. Overview

The addressable anti-tag capture agent collections, such as an positionally addressable array, contains a collection different capture agents, such as antibodies that bind to pre-selected and/or pre-designed polypeptide tags, such as epitope tags, with high affinity and specificity. A typical collection contains at least about 30, more prefereably 100, more preferably 500, most preferably at least 1000 capture agents, such as antibodies, that are addressable, such as by occupying a unique locus on an array or by virtue of being bound to bar-coded support, color-coded, or RF-tag labeled support or other such addressable format. Each locus or address contains a single type of capture agent, such as antibody, that binds to a single specific tag. Tagged proteins are contacted with the collection of receptors, such as antibodies in an array, under conditions

A21 cont.

suitable for complexation with the receptor, such as an antibody, via the epitope tag. As a result, proteins are sorted according to the tag each possesses.

Please replace the paragraph on page 35, lines 1-11, with the following paragraph:

3. Creating the master library for sorting

In this step, tags that encode each of the epitopes linked to each of the divider sequences are incorporated into the master library, which is typically a cDNA library. Any way known to those of skill in the art to add and incorporate a double stranded DNA fragment into nucleic acid may be used. In particular, a variety of ways are contemplated herein. These include (1) using PCR amplification to incorporate them (exemplified herein); (2) ligating them directly or via linkers (see below), the ligated product, if needed, can be amplified, and other methods described herein (see below) and that can be readily devised by those of skill in the art in light of the description herein.

A22

Please replace the paragraph on page 36, lines 1-15, with the following paragraph:

The master library is divided into pools, identified as D_1 - D_n, reacted with n number of addressable collections of antibodies, each collection containing antibodies with m different epitope specificities. Each collection, such as an array, is associated with one of the pools, such as by an optical code, including a bar code a notation or a symbol or a colored code, an electronic tag or other identifier, such as color or a identifiable chemical tag, on the collection or other such identifier. The reaction is performed under conditions whereby the epitopes bind to the antibodies specific therefor, and the resulting complexes of antibodies and epitope-tag-labeled molecules are screened using an assay that specifically identifies molecules that have a desired property. The particular collection(s) of antibodies and antibodies with a particular tag that includes molecules with the desired property are identified, thereby also identifying the

A23

A23 cont

particular D_n pool and epitope tag on the molecule, thereby reducing the diversity of the collection by n x m.

Please replace the paragraph beginning on page 36, line 25, through page 37, line 20, with the following paragraph:

a. Ligation to create circular plasmid vector for introduction of tags

As noted above, in addition to use of amplification protocols for introducing the primers into the library members, the primers may be introduced by direct ligation, such as by introduction into plasmid vectors that contain the nucleic acid that encode the tags and other desired sequences. Subcloning of a cDNA into double stranded plasmid vectors is well known to those skilled in the art. One method involves digesting purified double stranded plasmid with a site-specific restriction endonuclease to create 5' or 3' overhangs also known as sticky ends. The double stranded cDNA is digested with the same restriction endonuclease to generate complementary sticky ends. Alternately, blunt ends in both vector DNA and cDNA are created and used for ligation. The digested cDNA and plasmid DNA is mixed with a DNA ligase in an appropriate buffer (commonly, T4 DNA ligase and buffer obtained from New England Biolabs are used) and incubated at 16°C to allow ligation to proceed. A portion of the ligation reaction is transformed into E. coli that has been rendered competent for uptake of DNA by a variety of methods (electroporation, or heat shock of chemically competent cells are two common methods). Aliquots of the transformation mix are plated onto semi-solid media containing the antibiotic appropriate for the plasmid used. Only those bacteria receiving a circular plasmid gives rise to a colony on this selective media. Creation of a library of unique members is performed in a similar manner, however the cDNA being inserted into the vector is a mixture of different cDNA clones. These different cDNA clones are created via a wide variety of methods known to those skilled in the art.

A24

Please replace the paragraph on page 42, lines 6-29, with the following paragraph:

b. Ligation of sequences resulting in linear tagged cDNA

Following creation of the cDNA library, sequences are appended to cDNA clones via ligation. Linear, double stranded DNA containing each of the EDC sequence combinations is created via various methods (synthesis, digestion out of plasmid containing the sequences, assembly of shorter oligonucleotides, etc.). These linear dsDNAs containing the different EDC sequences, are mixed such that each individual is equally represented in the mixture. This mixture is combined with the double stranded cDNA library and ligated using a nucleic acid ligase in an appropriate buffer. This is preferably a DNA ligase, but an RNA ligase is used if the EDC tags are composed of RNA or are RNA/DNA hybrid molecules and the library is also in the form of an RNA or RNA/DNA hybrid. In one embodiment, the EDC sequence is blunt-ended on both ends yet only one end is phosphorylated such that ligation occurs in a directional manner (with respect to the EDC sequence), and the E sequence is brought into the same reading frame as the cDNA (at either the N or C terminus of the resulting protein). In another embodiment, the EDC sequence is blunt-ended at one end and has an overhang on the other end such that ligation occurs in a directional manner (see, Sambrook *et al.* (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Edition, Cold Spring Harbor Laboratory Press Chapter 8). The EDC sequences can be continuously double stranded, or partially double stranded with a single stranded central portion.

A25

Please replace the paragraphs beginning on page 50, line 14, through page 51, line 14, with the following paragraphs:

For these additional steps, the nucleic acid in the sample that contains the identified D_n is amplified with a set of primers that each contains a portion (designated FA_p) of each epitope-encoding tag (each designated E_p) sufficient to amplify the linked nucleic acid, but insufficient to reintroduce E_p, where each

A26

primer includes or is of a sequence of nucleotides of formula HO-FA-E_p, where p is an integer of 1 to m. This amplification introduces a different one of the epitope-encoding sequences into the nucleic acid to produce a collection of cDNA clones (a sublibrary of the original) that again contains all of the epitopes distributed among the sublibrary members.

In this second sorting step, if amplification is used to introduce the new set of tags, concatamer formation can be minimized by using a low concentration of the FA primers followed by an excess of primers encoding the common region, which region is introduced by the FA primer. After the FA primer is used, the common primers out compete the FA primers for incorporation, since the C region will then be incorporated into the template nucleic acid molecule.

A26 cont

Alternatively, as noted above, the new set of epitope-encoding sequences can be ligated via linkers to the template. To do this the template can be cut with a unique restriction enzyme and the linkers ligated. This can get rid of the existing epitope encoding nucleic acid and replace it with a new set of epitopes. Ligation can be followed by amplification with the common region. Other methods may also be used.

In creating the sublibrary for the second sorting step, as with the master library, it is necessary to use conditions that ensure that on the average each different molecule has a different tag and one of each kind is tagged. In this round, one tag, on the average, should attach to each of the different molecules. In this round, however, the diversity is much lower, since the first sorting step achieves an m x n reduction in diversity. Any of the methods described above to attach and distribute polypeptide tag-encoding sequences among the sublibrary members can be used.

Please replace the paragraphs beginning on page 51, line 24, through page 53, line 7, with th following paragraphs:

The nucleic acids in the resulting sublibrary are translated and the translated proteins contacted, such as under western blotting conditions, with one collection of capture agents (or a plurality of replicas thereof), such as antibodies, to form capture agent-protein complexes. The proteins in the complexes are screened to identify the capture agent, such as antibody or receptor, locus (or loci) that binds to the epitope linked to the protein of interest, thereby identifying the "E", the epitope sequence associated with the protein of interest. Nucleic acid molecules in the sublibrary that contain the identified "E", epitope sequence, designated E_q, are specifically amplified, with primers that include the formula 5′ FB_s 3′ (or 5′CFB_s3′), where each FB is sufficient to amplify the linked nucleic acid using an E_m portion of the epitope sequence and includes all or a portion of the E_m. This specifically amplifies the nucleic acid molecule of interest.

In summary, the diversity (Div) equals the total number of different molecules in a library (*i.e.*, 10^8), N = number of divisions D_1-D_n, which is the number of different collections of capture agents, such as 10^2; M = number of different epitope tags (and capture agents) E_1-E_m, such as 10^3. To start the method, a master tagged library is prepared, and divided N times. Portions of the N samples are translated and spotted onto N arrays each containing M capture agents (sort 1). At this stage M x N = 10^5. For the second sort, "M" new epitopes, such as 10^3 are used, the nucleic acid is translated and sorted onto one array of 10^3 capture agents, such as antibodies, thereby achieving a 10^8 reduction in diversity. As a result, each locus (or member of a collection if provided linked to particulate identifiable supports) in the array has a single type of protein as well as a single capture agent. The number of sorting steps can be any desired number, but is typically one or two. If a higher number of sorts are performed, then the sensitivity of the detection assay at the first sort should be very high, since, as a result of the diversity, the concentration of the protein of interest will be low. As noted above, M and N may be different each sorting step.

The process of nested sorting, which is applicable to sorting a variety of collections of molecules, particularly collections of proteins, DNA, small molecules and other collections is exemplified in Figures 1-18. The concept of nested sorting is illustrated in **Fig 1**. In this example, a master collection containing 74,088 different items, such as cDNA, is searched by randomly dividing the collection into 42 sublibraries (F1 sublibraries). After identifying which of the 42 F1 sublibraries contains the item of interest, such as by binding or reaction with a probe or by a protein-protein specific interaction, that group is further divided randomly into 42 new sublibraries (F2 sublibraries) and again the sublibrary containing the item of interest is identified. A final division of the F2 sublibrary containing the item of interest produces 42 new groups, each containing only one item. The item of interest can be uniquely identified based on its sorting lineage.

A27 cont.

Please replace the paragraph on page 53, lines 20-29, with the following paragraph:

Fig 3 illustrates the use of capture agent arrays, such as antibody arrays, as a tool for nested sorts of high diversity gene libraries. A master gene library is first randomly divided into a number of sublibraries by separate amplification, such as PCR, reactions. The amplification reactions use sets of unique sequences of nucleotides that encode preselected epitopes and incorporate these sequences into the genes by appropriate design of primers to specifically amplify different sublibraries of genes from the master template pool (F1 sublibraries). These amplification reactions are performed, for example, in 96-well (or 384-well or higher density) PCR plates with a compatible thermocycler.

A28

Please replace the paragraphs beginning on page 54, line 13, through page 55, line 4, with the following paragraphs:

The table in **Fig 3** illustrates how the number of initial divisions by PCR and the number of capture agents the array can be combined to search gene

A29

libraries containing, for example, from a million (10^6) to over a billion (10^9) different genes. For example, an initial gene library can be divided into 100 F1 sublibraries by amplification and then further divided using two arrays with capture agents recognizing 100 different epitopes. If the initial gene library contained 10^6 different genes, the F3 addresses in the sublibraries contain a single type of gene ($10^6/100/100/100 = 1$). An initial gene library divided into 1,000 F1 sublibraries by PCR amplification and then further divided using two arrays with capture agents recognizing 1,000 different epitopes to create the F2 and F3 sublibraries can be used to search 10^9 different genes ($10^9/1,000/1,000/1,000 = 1$).

Q29 con.

Dividing the gene libraries into sublibraries is based on the ability of a PCR amplification reaction to specifically amplify DNA sequences using pairs of primers. Although both primers need to hybridize to sequences on either end of the template DNA, a subset of template sequences can be amplified using a primer pair in which one of the primers is common to all of the template sequences and the other primer is specific for the gene sequence of interest. For example, specific genes are often amplified from cDNA libraries using one primer that is specific for the gene of interest and another that hybridizes to the oligo(dA) tail common to all of the cDNA molecules.

Please replace the paragraph beginning on page 55, line 25, through page 56, line 7, with the following paragraph:

Q30

In its simplest embodiment, consider an example of x tags at site X, y tags at site Y, and z tags at site Z. If these tags are used individually, then there are x + y + z combinations. If these tags are used in combination then there are (x)(y)(z) combinations. Assuming that the number of tags at each site (x, y and z) is one third the total (n), then for the case of individual use, $C = (n/3)x3 = n$ or there are as many total combinations (C) as there are tags; whereas for combinatorial use, there are $C = (n/3)^3$. As the number of individual tags at each site increases, the number of combinatorial tags increases at a

much higher rate (See Figure 19). With a greater number of effective tags, the number of members of the library per tag decreases. Fewer members per tag in the initial library results in either fewer sequential rounds of screening or lower numbers of clones that to be assessed with high throughput screening.

A30 cont.

Please replace the paragraphs beginning on page 56, line 23, through page 57, line 8, with the following paragraphs:

Recovery of a specific library pool with a combinatorial tag is done in substantially the way a library pool with a single tag is recovered. As described herein, one way to recover subpopulations from in the library is to use the polymerase chain reaction. For exemplification, assume that all three tags are at the C-terminus of an expressed protein such that the X tag is the most proximal to the library member, such as an scFv, followed by the Y tag and then the Z tag. The order of DNA segments on the coding strand of cDNA is:

A31

5′ Common>scFv>X>Y>Z 3′

A particular sub-population can be recovered by sequential rounds of PCR amplification starting with a common primer and a primer corresponding to the Z289 tag. The product from this reaction is used in the next reaction using the common primer and the Y132 tag primer. The product from this reaction is used in a subsequent reaction with the common primer and the X27 primer. After three sequential rounds of amplification, the products all correspond to library members, such as scFvs, that were originally tagged with the X27-Y132-Z289 combination.

Please replace the paragraph on page 57, lines 17-31, with the following paragraph:

One embodiment uses, for example, a single scaffold fusion protein containing multiple sites with inserted epitope tags. This spatially separates the epitopes and allows them all to be recognized without interference with one another. The following criteria are considered in selecting a protein scaffold: 1)

A32

known crystal structure to more easily identify surface exposed amino acids with high propensity for antigenicity, 2) free N and C-termini for fusion to the cDNA library of interest, 3) high levels of production and solubility in various protein expression systems (especially the E.coli periplasm), 4) capacity for *in vitro* transcription/translation, 5) absence of disulfide bonds, 6) wild-type protein is monomeric, 7) has capacity to increase solubility or function of scFvs. Using the crystal structure, positions are chosen for insertion of epitope tag libraries. These sites should be spatially separated epitopes that are relatively linear in nature (e.g. one side of an alpha helix, a turn between beta strands or a loop between helices).

A32 cont.

Please replace the paragraph beginning on page 58, line 10, through page 59, line 2, with the following paragraph:

Alternatively, each antibody can be identified, such as by linkage to optically encoded tags, including colored beads or bar coded beads or supports, or linked to electronic tags, such as by providing microreactors with electronic tags or bar coded supports (see, *e.g.*, U.S. Patent No. 6,025,129; U.S. Patent No. 6,017,496; U.S. Patent No. 5,972,639; U.S. Patent No. 5,961,923; U.S. Patent No. 5,925,562; U.S. Patent No. 5,874,214; U.S. Patent No. 5,751,629; U.S. Patent No. 5,741,462), or chemical tags (see, U.S. Patent No. 5,432,018; U.S. Patent No. 5,547,839) or colored tags or other such addressing methods that can be used in place of physically addressable arrays. For example, each antibody type can be bound to a support matrix associated with a color-coded tag (i.e. a colored sortable bead) or with an electronic tag, such as an radio-frequency tag (RF), such as IRORI MICROKANS® and MICROTUBES® microreactors (see, U.S. Patent No. 6,025,129; U.S. Patent No. 6,017,496; U.S. Patent No. 5,972,639; U.S. Patent No. 5,961,923; U.S. Patent No. 5,925,562; U.S. Patent No. 5,874,214; U.S. Patent No. 5,751,629; U.S. Patent No. 5,741,462; International PCT application No. WO98/31732; International PCT application No. WO98/15825; and, see, also U.S. Patent No.

A33

6,087,186). For the methods and collections provided herein, the antibodies of each type can be bound to the MICROKAN or MICROTUBE microreactor support matrix and the associate RF tag, bar code, color, colored bead or other identifier serves to identify the receptors, such as antibodies, and hence the epitope tag to which the receptor, such as an antibody, binds.

A33 cont.

Please replace the paragraph on page 59, lines 11-19, with the following paragraph:

Also, for exemplary purposes herein, reference is made to positional arrays. It is understood, however, that such other identifying methods can be readily adapted for use with the methods herein. It is only necessary that the identity (*i.e.*, epitope-tag specificity) of the receptor, such as an antibody, is known. The resulting collections of addressable receptors (*i.e.*, antibodies), whether in a two-dimensional or three-dimensional array, or linked to optically encoded beads or colored supports or RF tags or other format, can be employed in the methods herein.

A34

Please replace the paragraphs on page 60, lines 1-21, with the following paragraphs:

2. Preparation of the capture agents

The quality of the sorts is dependent on the quality of the collection of capture agents, such as antibodies, that make up the sorting array. In addition to requirements on binding affinity and specificity, the epitopes bound by the capture agents (antibodies) in the array determine the E, FA and FB sequences used as priming sites for the amplification reactions (PCRs). **Fig 12** outlines a high throughput screen for discovering immunoglobulin (Ig) produced from hybridoma cells for use in generating antibodies for use in the collections.

A35

Hybridoma cells are created either from non-immunized mice or mice immunized with a protein expressing a library of random disulfide-constrained heptameric epitopes or other random peptide libraries. Stable hybridoma cells

are initially screened for high Ig production and epitope binding. Ig production is measured in culture supernatants by ELISA assay using a goat anti-mouse IgG antibody. Epitope binding is also measured by ELISA assay in which the mixture of haptens (epitope tagged proteins) used for immunization are immobilized to the ELISA plate, and bound IgG from the culture supernatants is measured using a goat anti-mouse IgG antibody. Both assays are done in 96-well formats or other suitable formats. For example, approximately 10,000 hybridomas are selected from these screens.

A35 cont.

Please replace the paragraphs beginning on page 60, line 28, through page 63, line 8, with the following paragraphs:

The purified Ig are spotted separately onto a nitrocellulose filter using a standard pin-style arraying system. The purified Ig are also combined to produce a mixture with equal quantities of each Ig. The mixed Ig are bound to paramagnetic beads which are used as a solid-phase support to pan a library of bacteriophage expressing the random disulfide-constrained heptameric epitopes. The batch panning enriches the phage display library for phage expressing epitopes to the purified Ig. This enrichment dramatically reduces the diversity in the phage library.

A36

The enriched phage display library is then bound to the array of purified Ig and stringently washed. Ig-binding phage are detected by staining with an anti-phage antibody-HRP conjugate to produce a chemiluminescent signal detectable with a charge coupled device (CCD)-based imaging system. Spots in the array producing the strongest signals are cut out and the phage eluted and propagated. Epitopes expressed by the recovered phage are identified by DNA sequencing and further evaluated for affinity and specificity. This method generates a collection of high-affinity, high-specificity antibodies that recognize the cognate epitopes. Continued screening produces larger collections of antibodies of improved quality.

3. Preparation f anti-tag captur ag nt arrays

Each spot contains a multiplicity of capture agents, such as antibodies with a single specificity. Each spot is of a size suitable for detection. Spots on the order of 1 to 300 microns, typically 1 to 100, 1 to 50, and 1 to 10 microns, depending upon the size of the array, target molecules and other parameters. Generally the spots are 50 to 300 microns. In preparing the arrays, a sufficient amount is delivered to the surface to functionally cover it for detection of proteins having the desired properties. Generally the volume of antibody-containing mixture delivered for preparation of the arrays is a nanoliter volume (1 up to about 99 nanoliters) and is generally about a nanoliter or less, typically between about 50 and about 200 picoliters. This is very roughly about 10 million to 100,000 molecules per spot, where each spot has capture agents, such as antibodies, that recognize a single epitope. For example, if there are 10 million molecules and 1000 different ones in the protein mixture reacting with the locus, there are 10^4 of each type of molecule per spot. The size of the array and each spot should be such that positive reactions in the screening step can be imaged, preferably by imaging the entire array or a plurality therof, such as 24, 96, or more arrays, at the same time.

A136 cont.

A support (see below for exemplary supports), such as KODAK paper plus gelatin or other suitable matrix can be used, and then ink jet and stamping technology or other suitable dispensing methods and apparatus, are used to reproducibly print the arrays. The arrays are printed with, for example, a piezo or inkjet printer or other such nanoliter or smaller volume dispensing device. For example, arrays with 1000 spots can be printed. A plurality of replicate arrays, such as 24 or 48, 96 or more can be placed on a sheet the size of a conventional 96 well plate.

Among the embodiments contemplated herein, are sheets of arrays each with replicates of the antibody array. These are prepared using, for example, a piezo or inkjet dispensing system. A large number, for example, 1000 can be printed at a time using, for example a print head with 1000 different holes (like

a stamp with 500 μM holes). It can be fabricated from, for example, molded plastic with many holes, such as 1000 holes each filled with 1000 different capture agents, such as antibodies. Each hole can be linked to reservoirs that are linked to conduits of decreasing size, which ultimately dispense the capture agents, such as antibodies into the print head. Each array on the sheet can be spatially separated, and/or separated by a physical barrier, such as a plastic ridge, or a chemical barrier, such a hydrophobic barrier (*i.e.*, hydrogels separated by hydrophobic barriers). The sheets with the arrays can be conveniently the size of a 96 well plate or higher density. Each array contains a plurality of addressable anti-tag antibodies specific for the pre-selected set of epitope tags. For example, 33 x 33 arrays contain roughly 1000 antibodies, each spot on each array containing antibodies that specifically bind to a single pre-selected epitope. A plurality of arrays separated by barriers can be employed.

A36 cont.

For dispensing the antibodies onto the surface, the goal is functional surface coverage, such that a screened desired protein is detectable. To achieve this, for example, about 1 to 2 mg/ml from the starting collection are used and about 500 picoliters per antibody are deposited per spot on the array. The exact amount(s) can be empirically determined and depend upon several variables, such as the surface and the sensitivity of the detection methods. The antibodies are preferably covalently linked, such as by sulfhydryl linkages to amides on the surface.

Please replace the paragraphs beginning on page 63, line 19, through page 64, line 13, with the following paragraphs:

4. Preparation of other collections

A37

The capture agents are linked to beads or other particulate supports that are identifiable. For example, the capture agents are linked to optically encoded microspheres, such as those available from Luminex, Austin Tx, the contain fluorescent dyes encapsulated therein. The microsphere, which encapsulate dyes, are prepared from any suitable material (see, *e.g.*, International PCT

application Nos. WO 01/13119 and WO 99/19515; see description below), including styrene-ethylene-butylene-styrene block copolymers, homopolymers, gelatin, polystyrene, polycarbonate, polyethylene, polypropylene, resins, glass, and any other suitable support (matrix material), and are of a size of a about a nanometer to about 10 millimeters in diameter. By virtue of the combination of, for example two different dyes at ten different concentrations, a plurality microspheres (100 in this instance), each identifiable by a unique fluorescence, are produced.

A37 Cont.

Alternatively, combinations of chromophores or colored dyes or other colored substances are encapsulated to produce a variety of different colors encapsulated in microspheres or other particles, which are then used as supports for the capture agents, such as antibodies. Each capture agent, such as an antibody, is linked to a particular colored bead, and, is thereby identifiable. After producing the beads with linked capture agents, such as antibodies, reaction with the epitope-tagged molecules can be performed in liquid phase. The beads that react with the epitopes are identified, and as a result of the color of the bead the particular epitope and is then known. The sublibrary from which the linked molecule is derived is then identified.

Please replace the paragraphs beginning on page 64, line 25, through page 66, line 5, with the following paragraphs:

Supports that are also contemplated for use herein include fluophore-containing or -impregnated supports, such as microplates and beads (commercially available, for example, from Amersham, Arlington Heights, IL; plastic scintillation beads from Nuclear Technology, Inc., San Carlos, CA and Packard, Meriden, CT, and colored bead-based supports (fluorescent particles encapsulated in microspheres) from Luminex Corporation, Austin, TX (see, International PCT application No. WO/0114589, which is based on U.S. application Serial No. 09/147,710; see International PCT application No. WO/0113119, which is U.S. application Serial No. 09/022,537). The

A38

microspheres from Luminex, for example, are internally color-coded by virtue of the encapsulation of fluorescent particles and can be provided as a liquid array. The capture agents, such as antibodies (epitopes) are linked directly or indirectly by any suitable method and linkage or interaction to the surface of the bead and bound proteins can be identified by virtue of the color of the bead to which they are linked. Detection can be effected by any means, and can be combined with chromogenic or fluorescent detectors or reporters that result in a detectable change in the color of the microsphere (bead) by virtue of the colored reaction and color of the bead. For the bead-based arrays, the anti-tag capture agents are attached to the color-coded beads in separate reactions. The code of the bead identifies the capture agent, such as antibody, attached to it. The beads can then be mixed and subsequent binding steps performed in solution. They can then be arrayed, for example, by packing them into a microfabricated flow chamber, with a transparent lid, that permits only a single layer of beads to form resulting in a two-dimensional array. The beads on which a protein is bound identified, thereby identifying the capture agent and the tag. The beads are imaged, for example, with a CCD camera to identify beads that have reacted. The codes of the such beads are identified, thereby identifying the capture agent, which in turn identifies the polypeptide tag and, ultimately, the protein of interest.

Q38 cont.

The support may also be a relatively inert polymer, which can be grafted by ionizing radiation to permit attachment of a coating of polystyrene or other such polymer that can be derivatized and used as a support. Radiation grafting of monomers allows a diversity of surface characteristics to be generated on supports (see, *e.g.*, Maeji *et al.* (1994) *Reactive Polymers 22*:203-212; and Berg *et al.* (1989) *J. Am. Chem. Soc. 111*:8024-8026). For example, radiolytic grafting of monomers, such as vinyl monomers, or mixtures of monomers, to polymers, such as polyethylene and polypropylene, produce composites that have a wide variety of surface characteristics. These methods have been used

A38 Cont.—

to graft polymers to insoluble supports for synthesis of peptides and other molecules.

Please replace the paragraph beginning on page 68, line 26, through page 69, line 7, with the following paragraph:

For example, U.S. Patent No. 5,403,750, describes the preparation of polyurethane-based polymers. U.S. Pat. No. 4,241,537 describes a plant growth medium containing a hydrophilic polyurethane gel composition prepared from chain-extended polyols; random copolymerization can be performed with up to 50% propylene oxide units so that the prepolymer is a liquid at room temperature. U.S. Pat. No. 3,939,123 describes lightly crosslinked polyurethane polymers of isocyanate terminated prepolymers containing poly(ethyleneoxy) glycols with up to 35% of a poly(propyleneoxy) glycol or a poly(butyleneoxy) glycol. In producing these polymers, an organic polyamine is used as a crosslinking agent. Other supports and preparation thereof are described in U.S. Patent Nos. 4,177,038, 4,175,183, 4,439,585, 4,485,227, 4,569,981, 5,092,992, 5,334,640, 5,328,603.

A39

Please replace the paragraph on page 69, lines 18-27, with the following paragraph:

U.S. Patent No. 4,171,412 describes specific supports based on hydrophilic polymeric gels, preferably of a macroporous character, which carry covalently bonded D-amino acids or peptides that contain D-amino acid units. The basic support is prepared by copolymerization of hydroxyalkyl esters or hydroxyalkylamides of acrylic and methacrylic acid with crosslinking acrylate or methacrylate comonomers are modified by the reaction with diamines, aminoacids or dicarboxylic acids and the resulting carboxyterminal or aminoterminal groups are condensed with D-analogs of aminoacids or peptides. The peptide containing D-aminoacids also can be synthesized stepwise on the surface of the carrier.

A40

Please replace the paragraphs beginning on page 74, line 14, through page 75, line 24, with the following paragraphs:

F. Use of the methods for identification of proteins of desired properties from a library

1. Arraying capture agents

Q41

The capture agent molecules to which the epitope tags specifically bind are linked to supports, such as identifiable beads, such as microsheres, or solid surfaces. Linkage can be effected through any suitable bond, such as ionic, covalent, physical, van der Waals bonds. It can be effected directly or via a suitable linker. For exemplary purposes arraying on surfaces is described.

Purified antibodies (1 μl at a concentration of 1-2 mg/ml in a buffer of 0.1 M PBS (phospahte buffered saline, pH 7.4) on glycerol (1-20% vol/vol), are spotted onto a membranes (such as; UltraBind membrane, Pall Gelman; FAST nitrocellulose coated slides, Schleicher & Schuell), chemically deactivated glass slides, superaldehyde slides (Telechem), polylysine coated glass, activated glass, or specific thin films and self-assembled monolayers International PCT application Nos WO 00/04389, WO 00/04382 and WO 00/04390), using an automated arraying tool (such as systems available from, for example, Microsys; PixSys NQ; Cartesian Technologies; BioChip Arrayer; Packard Instrument Company; Total Array System; BioRobotics; Affymetrix 417 Arrayer; Affymetrix, and others). The spots are allowed to air dry for a suitable period of time, 1-2 minutes or more, typically 30 min to 1 h. Two membrane attachments are described. The UltraBind membrane (Pall Gelman) contains active aldehyde groups that react with primary amines to form a covalent linkage between the membrane and the capture agent, such as an antibody. Unreacted aldehydes are blocked by incubation with suitable blocking solution, such as a solution of 50 mM PBS, pH 7.4, 2 % bovine serum albumin (BSA) or with BBSA-T (a protein-containing solution such as Blocker BSA™ (Pierce) diluted to 1x in phosphate-buffered saline (PBS) with Tween-20 (polyoxyethylenesorbitan monolaurate;

Sigma) added to a final concentration of 0.05% (vol:vol)) for a suitable time, such as about 30 minutes. The filter can be rinsed with PBS.

Capture agents, such as antibodies, also can be deposited onto membranes, such as, for example, nitrocellulose paper (Schliecher& Schuell) with, for example, an inject printer (*i.e.*, Canon model BJC 8200, color inject printer), modified for this use and connected to a computer, such as a personal computer (PC). Such modifications, include, removal of the color ink cartridges from the print head and replacement with, for example, 1 milliliter pipette tips, which are hand-cut to fit in a sealed manner over the inkpad reservoir wells in the print head. Antibody solutions are pipetted into the pipette tips reservoirs that are seated on the inkpad reservoirs.

Please replace the paragraphs beginning on page 75, line 30, through page 77, line 16, with the following paragraphs:

Purified capture agents, such as antibodies can also be spotted onto FAST nitrocellulose coated slides, (Schleicher & Schuell). Nitrocellulose binds proteins by noncovalent adsorption. Nitrocellulose binds approximately 100 μg per cm^2. After binding of the capture agents, such as antibodies, remaining binding sites are blocked by incubation with a solution of 50 mM PBS, pH 7.4, 2 % bovine serum albumin (BSA) or BBSA-T for a suitable time, such as for 30 minutes.

Direct binding of antibodies to the nitrocellulose results in non-oriented binding. The percentage of active immobilized antibody molecules can be increased by binding to nitrocellulose that has been coated with an antibody capture protein (such as protein A, protein G or anti-IgG monoclonal antibody). The antibody capture proteins are bound to the nitrocellulose before application of the library proteins, such as tagged antibodies, with an arrayer. Biotinylated antibodies can also be printed onto surfaces coated with avidin or strepavidin. The size and spacing of the spots can be adjusted depending on the filter used

and the sensitivity of the assay. Typical spots are about 300-500 μm in diameter with 500-800 μm pitch.

Antibodies can also be printed onto activated glass substrates. Prior to printing the glass is cleaned ultrasonically in succession with a 1:10 dilution of detergent in warm tap water for 5 minutes in Aquasonic Cleaning Solution (VWR), multiple rinses in distilled water and 100% methanol (HPLC grade) followed by drying in a class 100 oven at 45° C. Clean glass is chemically functionalized by immersion in a solution of 3-aminopropyltriethoxysilane (APTS) (5% vol/vol in absolute ethanol) for 10 minutes. The glass is then rinsed in 95% ethanol, allowed to air dry, and then heated to 80° C in a vacuum oven for 2 hours to cure. The surface can then be further modified to bind primary amines or free sulfhydryl groups in the antibody or avidin or strepavidin linked to the antibody with biotin. To create an amine-reactive surface, the functionalized glass is treated with a solution of *Bis*[sulfosuccinimidyl]suberate (BS^3)(5 mg/ml in PBS, pH 7.4) for 20 minutes at room temperature. The *N*-hydroxysuccinimide (NHS)-activated glass surface is rinsed with distilled water and placed in a 37° C dust-free class 100 oven for 15 minutes to dry. Antibodies can be directly attached to this surface or the surface can be coated with a protein such as protein A that binds the antibodies, protein G or anti-IgG monoclonal antibody or avidin/strepavidin, to bind biotinylated proteins. To create a sulfhydryl-reactive surface, the functionalized glass is treated with a solution of sulfosuccinimidyl 4-[*N*-maleimidomethyl]-cyclohexane-1-carboxylate (Sulfo-SMCC) for 20 minutes at room temperature. The maleimide-activated glass surface is rinsed with distilled water and placed in a 37° C dust-free class 100 oven for 15 minutes to dry. To create a biotinylated surface, the functionalized glass is treated with a solution of EZ-link Sulfo-NHS-LC-Biotin (Pierce) for 20 minutes at room temperature. The biotinylated glass surface is rinsed with distilled water and placed in a 37° C dust-free class 100 oven for 15 minutes to dry. The same immobilization strategies described above also can be used in self-assembled monolayers formed on top of inorganic thin films.

A42 cont.

Please replace the paragraph beginning on page 77, line 30, through page 78, line 21, with the following paragraph:

A43

Initially a set of oligonucleotides containing various functional domains are added to the 3' ends of a gene to be mutated by incorporation of a primer that contains sequences of nucleotides that hybridize to the gene and also additional sets of sequences, (designated E for "Epitopes" D for "Divider", and C for "Common"). The E D C sequences constitute sets of sequences, each defined by the functions in the nucleic acid. As noted, the E sequences encode the epitopes specifically recognized by antibodies in the collection. They are incorporated in-frame with the coding sequences of the gene to be mutated and are expressed as a fusion with the parent protein. The D sequences are unique sequence sets downstream from the epitopes. They serve as specific priming sites to "Divide" the master group. They can be non-coding sequences and do not necessarily end up being part of the expressed mutated proteins. The C sequence is a sequence "Common" to all of the genes and provides a means for simultaneous PCR amplification of all the gene templates. As noted previously, in certain embodiments the D and/or C sequences are optional. Importantly, the E and D sequences are randomly distributed among the resulting DNA molecules. For example, 100 E sequences and 100 D sequences combine to create 10,000 (100 x 100 = 10,000) uniquely tagged cDNA molecules. Likewise, 1,000 E sequences and 1,000 D sequences combine to create 1,000,000 (1,000 x 1,000 = 1,000,000) uniquely tagged cDNA molecules.

Please replace the paragraph on page 81, lines 5-19, with the following paragraph:

A44

The amplified genes from the F2 sublibrary are expressed *in vitro*, incubated with the antibody array, re-probed and analyzed. As before, "bright spots" in this array identifies the E sequence associated with the mutant gene of interest. At this point in the sort, the gene of interest (as illustrated in **Fig 4**) is known to be in the $F1_{50}$ and $F2_{23}$ sublibraries and contains the E45 sequence

($F1_{50}/F2_{23}/F3_{45}$). This information identifies a specific gene that can be amplified using a primer specific for the E45 sequence (FB_{45} C). The FB C primer is composed of two functional regions. The FB region contains sequences corresponding to a downstream fragment (Fragment B) of the E sequence present in the template. FB can contain all or part of E; C is optional. FB contains any part, up to and including all of the E encoding sequence, to confer hybridization specificity. As before, the C region encodes a common sequence for amplification. The resulting amplified genes represent an F3 sublibrary ($F3_{45}$).

a44 cont.

Please replace the paragraph on page 82, lines 9-15, with the following paragraph:

The method of recombinant antibody library construction can be adapted for use with the sorting methods herein. This is accomplished by incorporating the E D C sequences into the V_L chain genes before assembly with the V_H chain and linker sequences. After the recombinant antibody library has been tagged with the E D C sequences, it is sorted by division into the F1 sublibraries followed by screening with the arrays as described above.

a45

Please replace the paragraph beginning on page 82, line 22, through page 83, line 2, with the following paragraph:

Fig 6 illustrates how E D C sequences are put onto the V_L chain genes by primer incorporation. The V_H chain genes are cloned using standard methods. The mRNA isolated from spleen cells or PBLs is converted to cDNA using a universal oligo dT primer or IG gene-specific primers. The V_H genes are then specifically amplified using a set of primers that are complementary to consensus sequences that flank these genes. The V_{HBACK} primer also contains promoter sequences that are required for *in vitro* transcription and translation of the assembled gene and/or allows subcloning into plasmid vectors for *in vivo* expression in cells, such as, but are not limited to, bacterial, yeast, insect and mammalian cells.

a46

Please replace the paragraphs beginning on page 83, line 25, through page 84, line 15, with the following paragraphs:

Fig 8 outlines a method for searching a recombinant antibody library. The V_H and V_L genes are cloned as described above and the E D C sequences are added to the 3'-end of the antibody genes to create the master library. The F1 sublibraries are created using the D C set of PCR primers. The illustration depicts 100 F1 sublibraries, shows D C primers for $F1_2$, $F1_{50}$ and $F1_{99}$, and shows the amplified product from the $F1_{50}$ reaction.

Transcription and translation of the $F1_{50}$ sublibrary genes produces a variety of recombinant capture agents, such as antibodies, that can be randomly grouped according to the epitopes (E sequences) they contain. The expressed proteins are bathed over the array and allowed to sort onto spots in the array that contain antibodies that bind their specific epitope tags. After the scFvs from sublibrary $F1_{50}$ are bound to the array, labeled antigen is bathed over the array. The label on the antigen can be a chemical tag, such as biotin, used to bind a secondary detection reagent such as strepavidin conjugated HRP, or the antigen can be epitope tagged and detection achieved with an anti-epitope antibody-HRP complex. After binding, the array is washed, probed, and analyzed. Analysis is typically by photon collection using a CCD-based imaging detector and photons are typically produced by local enzymatic chemiluminescent reactions. Again, the "brightest spot" contains the recombinant antibody with the greatest affinity having bound the greatest amount of antigen.

Please replace the paragraphs beginning on page 84, line 30, through page 85, line 27, with the following paragraphs:

The amplified genes from the F2 sublibrary are expressed *in vitro* or in *in vivo* systems, incubated with the antibody array, re-probed and analyzed. As previously, "bright spots" in this array identifies the E sequence associated with the recombinant antibody gene of interest. At this point in the sort, the gene of

interest (as illustrated in **Fig 8**) is known to be in the $F1_{50}$ and $F2_{23}$ sublibraries and contains the E45 sequence ($F1_{50}/F2_{23}/F3_{45}$). This information identifies a specific gene that can be amplified using a primer specific for the E45 sequence (FB_{45} C). The resulting amplified genes represent an F3 sublibrary ($F3_{45}77$) that contains a single type of recombinant antibody.

H. Detection of bound antigen(s)

Bound polypeptide-tagged molecules can be detected by any suitable method known to those of skill in the art and is a function of the target molecules. Exemplary detection methods include the use of chemiluminescence and bioluminescence generating reagents, such as horse radish peroxidase (HRP) systems and luciferin/luciferase systems, alkaline phosphatase (AP), labeled antibodies, fluorophores and isotopes. These can be detected using film, photon collection, scanning lasers, waveguides, ellipsometry, CCDs and other imaging means.

Q48 Cont.

As noted, uses of the addressable anti-tag capture agent collections include, but are not limited to: searching a recombinant antibody scFv library to identify scFV including, but is not limited to, finding single antigen or multiple antigens; searching mutation libraries, including tagging mutant libraries; mutation by error prone PCR; mutation by gene shuffling for searching for small molecule binders, searching for increased antibody affinity, searching for enhanced enzymatic properties (AP, HRP, Luciferase, GFP); searching for sequence-specific DNA binding proteins; searching a cDNA library for protein-protein interactions; and any other such application.

Please replace the paragraph beginning on page 86, line 26, through page 87, line 14, with the following paragraph:

Q49

The purified antibodies are arrayed by robotic spotting onto a filter and are also separately mixed then bound to paramagnetic beads to create a substrate for panning high affinity epitopes from a filamentous M13 bacteriophage library displaying random cysteine-constrained heptameric amino

acid sequences. The phage library is enriched for phage displaying high affinity epitopes by mixing the phage library with the antibody-coated beads and washing away loosely-bound phage from the beads ("panning"). Several rounds of panning leads to a highly enriched library containing phage that tightly bind to the monoclonal antibodies present in the collection. To separate and identify high affinity phage-antibody pairs, the enriched phage library is incubated with the filter containing the arrayed antibodies under high stringency binding conditions. Phage bound to antibodies on the filter are identified by staining with HRP-conjugated anti-phage antibodies and a chemiluminescent substrate to produce a luminescent signal. The signal is quantified using a high resolution CCD camera imaging device. High affinity binding phage are recovered from the filter and propagated. Several independent phage clones recovered from each spot are sequenced to identify consensus high-affinity epitopes for the corresponding antibodies.

Q49 cont.

Please replace the paragraph beginning on page 88, line 27, through page 89, line 21, with the following paragraph:

b. Isolating hybridoma cells

Stable hybridomas are selected by growth for several days in poor medium. The medium is then replaced with fresh medium and single hybridomas are isolated by limited dilution cloning. Because hybridoma cells have a very low plating efficiency, single cell cloning is done in the presence of feeder cells or conditioned medium. Freshly isolated spleen cells can be used as feeder cells as they do not grow in normal tissue culture conditions and are lost during expansion of the hybridoma cells. In this procedure a spleen is aspectically removed from a mouse and disrupted. Released cells are washed repeatedly in medium containing 10% FBS. A spleen typically produces 100 ml of 10^6 cells per ml. The feeder cells are plated in 96-well plates, 50 μl per well, and grown for 24 h. Healthy hybridoma cells are diluted in medium containing 20% FBS, 2 x OPI to a concentration of 20 cells per ml. Cells should be as free of clumps as

Q50

possible. Add 50 µl of the diluted hybridoma cells to the feeder cells, final volume is 100 µl. Clones begin to appear in 4 days. Alternatively single cells can be isolated by single-cell picking by individually pipetting single cells and then depositing in wells containing feeder cells. Single cells can also be obtained by growth in soft agar. Once healthy, stable cultures are achieved the cells are maintained by growth in DME (or RPMI 1640) medium supplemented with 10% FBS. Stable cells can be stored in liquid nitrogen by slow freezing in medium containing a cryoprotectant such as dimethylsulfoxide (DMSO). The amount of antibody being produced by the cells is determined by measuring the amount of antibody in the culture supernatants by the ELISA method.

A50 cont.

Please replace the paragraph on page 91, lines 7-17, with the following paragraph:

4. Panning a phage display library on paramagnetic beads

A51

A phage library containing random cysteine-constrained peptides expressed as part of an N-terminal genetic fusion to the gene III protein (gIII) of the filamentous bacteriophage M13 is constructed essentially as described (Kay *et al.* (1996) *Phage Display of Peptides and Proteins: A Laboratory Manual*, Academic Press, San Diego). The random peptides are encoded by a DNA insert assembled from synthetic degenerate oligonucleotides and cloned into gIII. These libraries are available commercially (Ph.D.-C7C™ Disulfide Constrained Peptide Library Kit, New England Biolabs). The Ph.D.-C7C™ library contains approximately 3.7×10^9 independent clones.

Please replace the paragraphs beginning on page 91, line 25, through page 93, line 3, with the following paragraphs:

A52

Combine the phage-antibody solution with Dynabeads Pan Mouse IgG (Dynal). The beads are supplied as a suspension in PBS, pH 7.4, 0.1% BSA, 0.02% sodium azide. The beads are washed with TBS (50 mM Tris-HCl (pH 7.4), 150 mM NaCl) several times prior to mixing with phage. The beads are

separated from the solution by application of a magnet (Magnetic Particle Concentrator, Dynal). Add the phage-antibody solution to a concentration of 0.1 $\mu g/10^7$ beads and incubate at 4°C for 30 minutes with gentle tilting and rotation. Inclusion of the human antibody prevents selection of phage that bind to the human antibody immobilized on the Dynabeads. Additionally, inclusion of human proteins from a lysed human cell as a blocker will prevent the selection of phage epitopes also present in human cells. The selected antibody-phage pairs should not be competed with proteins naturally present in the samples to be tested.

In the next step of the method, remove the fluid using the magnet and resuspend the beads in a Wash buffer of 1 ml of TBST. Repeat wash step 10 times. After the last wash step, elute the captured phage by suspending the beads in 1 ml of 0.2 M glycine-HCl, pH 2.2, 1 mg/ml BSA and incubating for 10 minutes at room temperature before recovering the fluid. The pH of the recovered fluid is immediately neutralized with the addition of 0.15 ml of 1 M Tris, pH 9.1. A small aliquot of the eluate is titered by infecting ER2738 *Escherichia coli* (*E. coli*) cells on LB-Tet plates.

Q52 CONT.

Amplify the eluate by the addition of 20 ml of a mid-log culture of ER2738 *E. coli* and continue to grow in LB-Tet for 4.5 hours. Separate phage virions from *E. coli* cells by centrifugation at 10,000 rpm, 10 minutes, and transfer to fresh tube. Repeat, transferring the upper 80% of the supernatant to a fresh tube. Concentrate the phage by the addition of 1/6 volume of PEG/NaCl (20% w/v polyethylene glycol-8000, 2.5 M NaCl) followed by precipitation overnight at 4°C. The phage are recovered by centrifugation at 10,000 rpm for 15 minutes and the pellet is resuspended in 1 ml of TBS. Re-precipitate the phage in a microcentrifuge tube with PEG/NaCl and resuspend the pellet in 0.2 ml TBS, 0.02% sodium azide. Microcentrifuge for 1 minute to remove any residual material. The supernatant is the amplified eluate. Titer the amplified eluate and repeat the panning as described above 3 times. With each round of panning and amplification, the pool of phage becomes enriched for phage that

Q52 cont.

bind the antibodies. If the concentration of phage used as input is kept constant, an increase in the number of phage recovered should occur. Phage can be stored at 4°C or diluted 1:1 with sterile glycerol and stored at -20°C.

Please replace the paragraph beginning on page 93, line 24, through page 94, line 13, with the following paragraph:

6. Recovery of phage from filter and sequencing the epitopes

Phage can be recovered from the filter by cutting out the spots containing phage identified from the imaging. Phage are eluted from the filter by suspending the filter piece in 0.5 ml of 0.2 M glycine-HCl, pH 2.2, 1 mg/ml BSA and incubating for 10 minutes at room temperature before recovering the fluid. The pH of the recovered fluid is immediately neutralized with the addition of 0.075 ml of 1 M Tris, pH 9.1. A small aliquot of the eluate is titered by infecting ER2738 *E. coli* cells on LB-Tet plates. Isolated plaques (typically 10 plaques) are picked for DNA isolation and sequenced to define a consensus epitope. Plaques are amplified by inoculating 1 ml cultures of ER2738 *E. coli* cells freshly diluted 1:100 from a healthy mid-log culture, using a sterile pipet tip or toothpick and incubated at 37°C for 4 to 5 hours with shaking. Phage are recovered by microcentrifugation for 30 seconds, and 0.5 ml of the supernatant transferred to a fresh tube and 0.2 ml of PEG/NaCl is added and allowed to stand at room temperature after gentle mixing for 10 minutes. Pellet the phage by centrifugation for 10 minutes at top speed in a microcentrifuge. Discard any remaining supernatant and thoroughly suspend the pellet in 0.1 ml iodine buffer and 0.25 ml ethanol to precipitate single-stranded DNA. The DNA pellets are washed in 70% ethanol and air-dried. DNA is sequenced by standard methods.

Q53

Please replace the paragraphs beginning on page 95, line 26, through page 96, line 21, with the following paragraphs:

Q54

Recombinant antibodies are expressed and purified as described (McCafferty *et al.* (1996) *Antibody engineering: A practical Approach*, Oxford

University Press, Oxford). Briefly, the gene encoding the recombinant antibody is cloned into an expression plasmid containing an inducible promoter. The production of an active recombinant antibody is dependent on the formation of a number of intramolecular disulfide bonds. The environment of the bacterial cytoplasm is reducing, thus preventing disulfide bond formation. One solution to this problem is to genetically fuse a secretion signal peptide onto the antibody which directs its transport to the non-reducing environment of the periplasm (Hanes *et al.* (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*:4937-4942).

Alternatively, the antibodies can be expressed as insoluble inclusion bodies and then refolded *in vitro* under conditions that promote the formation of the disulfide bonds. Inoculate 0.5 liters of LB medium containing an appropriate antibiotic and shake for 10 hours at 32 °C. Use the starter culture to inoculate 9.5 liters of production medium (3 g ammonium sulfate, 2.5 g potassium phosphate, 30 g casein, 0.25 g magnesium sulfate, 0.1 mg calcium chloride, 10 ml M-63 salts concentrate, 0.2 ml MAZU 204 Antifoam (Mazer Chemicals), 30 g glucose, 0.1 mg biotin, 1 mg nicotinamide, appropriate antibiotic, per liter, pH 7.4). Ferment using a Chemap (or like) fermenter at pH 7.2, aeration at 1:1 v/v Air to medium per minute, 800 rpm agitation, 32° C. When the absorbance at 600 nm reaches 18-20, raise temperature to 42° C for 1 hour then cool to 10° C for 10 minutes before harvesting cell paste by centrifugation at 7,000 x g for 10 minutes. Recovery is typically 200-300 g wet cell paste from a 10 liter fermentation and should be kept frozen.

Please replace the paragraph beginning on page 97, line 30, through page 98, line 16, with the following paragraph:

After depositing the capture antibodies, including anti-HA tag capture antibodies onto a membrane, such as a nitrocellulose membrane, they are dried at ambient temperature and relative humidity for a suitable time period (*e.g.*, 10 minutes to 3 h, which can be determined empirically). After drying, membranes with deposited and dried anti-HA capture antibodies are blocked, if necessary,

with a protein-containing solution such as Blocker BSA™" (Pierce) diluted to 1x in phosphate-buffered saline (PBS) with Tween-20 (polyoxyethylenesorbitan monolaurate; Sigma) added to a final concentration of 0.05% (vol:vol) to eliminate background signal generated by non-specific protein binding to the membrane. For subsequent description contained herein, blocking agent is referred to as BBSA-T, and PBS with 0.05% (vol:vol) Tween-20 is referred to as PBS-T. Blocking times can be varied from 30 mm to 3 h, for example. For all subsequent incubations (except for washes) described below for this procedure, incubation times are varied from about 20 min to 2 h. Likewise, incubation temperatures can be varied from ambient temperature to about 37° C. In all instances, the precise conditions can be determined empirically.

Q55 Cont

Please replace the paragraph beginning on page 98, line 30, through page 99, line 17, with the following paragraph:

Membranes with deposited anti-HA capture antibodies and bound HA-HFN scFcv antigen is then washed a plurality of times (typically 3 times) with PBS-T, for suitable times (typically 3 to 5 min per wash, for example), at various temperature. Membranes with deposited anti-HA capture antibodies and bound HA-HFN scFv are then incubated with, for purposes of demonstration, biotinylated human fibronectin (Bio-HFN), which is an antigen that will be recognized by the capture HA-HFN scFv. Bio-HFN is serially diluted (*e.g.*, from 1 to 10 μg/ml) in BBSA-T. The resulting membranes are washed a suitable number of time (typically 3) with PBS-T for a suitable period of time (typically 3 to 5 min per wash) at various temperatures, and are then incubated with Neutravidin·HRPO (Pierce) serially diluted (*e.g.*, 1:1000 to 1:100,000 in BBSA-T). The resulting membranes are washed as before, rinsed with PBS and developed with Supersignal™ ELISA Femto Stable Peroxide Solution and Supersignal™ ELISA Femto Lumino Enhancer Solution (Pierce), and then imaged using an imaging system, such as, for example, a Kodak Image Station 440CF

Q56

or other such imaging system. A 1:1 mixture of peroxide solution:luminol is prepared and a small volume is plated on the platen of the image station.

A54 Cont.

Please replace the paragraph beginning on page 101, line 17, through page 102, line 9, with the following paragraph:

The cDNA library construction is done essentially as described (Sambrook *et al.* (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Edition, Cold Spring Harbor Laboratory Press). First strand synthesis is done by mixing the following at 4° C to 50 μl final volume: 10 μg mRNA (poly(A)$^+$RNA), 10 μg of V_{LFOR}-common primer mix (V_{LFOR}-common is described below), 50 mM Tris-HCl, pH 7.6, 70 mM potassium chloride, 10 mM magnesium chloride, dNTP mix (1 mM each), 4 mM dithiothreitol, 25 units RNase inhibitor, 60 units murine reverse transcriptase (Pharmacia). Incubate for 1 hour at 37° C. For the second strand synthesis a mixture of the following is directly added to the first strand synthesis solution to a final volume of 142 μl: 5 mM magnesium chloride, 70 mM Tris-HCl, pH 7.4, 10 mM ammonium sulfate, 1 unit RNAse H, 45 units *E. coli* DNA polymerase I, and allowed to incubate at room temperature for 15 minutes. To this mix is added 5 μl of 0.5 M EDTA, pH 8.0, to stop the reaction. The final volume should be 150 μl. The newly synthesized cDNA is purified by extraction with an equal volume of phenol:chloroform and the unincorporated dNTPs are separated by chromatography through Sephadex G-50 equilibrated in TE buffer (10 mM Tris-HCl, 1 mM EDTA), pH 7.6, containing 10 mM sodium chloride. The eluted DNA is precipitated by the addition of 0.1 x volume 3 M sodium acetate (pH 5.2) and 2 volumes of ethanol incubated at 25 °C for at least 15 minutes and recovered by centrifugation at 12,000g for 15 minutes at 4 °C, washed with 70% ethanol, air dried, then redissolved in 80 μl of TE (pH 7.6).

A57

Please replace the paragraph beginning n pag 102, lin 30, through pag 104, line 17, with th f llowing paragraph:

The cDNA library construction is done essentially as described (Sambrook *et al.* (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Edition, Cold Spring Harbor Laboratory Press). First strand synthesis is done by mixing the following at 4° C to 50 μl final volume; 10 μg mRNA (poly(A)$^+$RNA), 10 μg of 5'-restriction sequence-oligo(dT)$_{12-18}$ primers, 50 mM Tris-HCl, pH 7.6, 70 mM potassium chloride, 10 mM magnesium chloride, dNTP mix (1 mM each), 4 mM dithiothreitol, 25 units RNase inhibitor, 60 units murine reverse transcriptase (Pharmacia). Incubate for 1 hour at 37° C. For the second strand synthesis, a mixture of the following is directly added to the first strand synthesis solution to a final volume of 142 μl: 5 mM magnesium chloride, 70 mM Tris-HCl, pH 7.4, 10 mM ammonium sulfate, 1 unit RNAse H, 45 units *E. coli* DNA polymerase I, 1 U of the restriction enzyme recognizing the site on the 5'-end of the oligo (dT) primer and allowed to incubate at room temperature for 15 minutes. To this mix is added 5 μl of 0.5 M EDTA, pH 8.0, to stop the reaction. The final volume should be 150 μl. The newly synthesized cDNA is purified by extraction with an equal volume of phenol:chloroform and the unincorporated dNTPs are separated by chromatography through Sephadex G-50 equilibrated in TE buffer (10 mM Tris-HCl, 1 mM EDTA), pH 7.6, containing 10 mM sodium chloride. The eluted DNA is precipitated by the addition of 0.1 x volume 3 M sodium acetate (pH 5.2) and 2 volumes of ethanol incubated at 25 °C for at least 15 minutes and recovered by centrifugation at 12,000g for 15 minutes at 4 °C, washed with 70% ethanol, air dried, then redissolved in 80 μl of TE (pH 7.6) and the DNA concentration measured by absorbtion at 260 nm. The cDNA library is then tagged by the addition of unique linkers to the restriction digested 3'-end of the cDNA molecules. Linkers are prepared as described below and ligated to the purified cDNA in a reaction containing an equal number of cDNA and linker molecules, 10 U T4 DNA ligase (100 U/μl), 1 μl 10 mM ATP, 1 μl Ligation buffer (0.5 M Tris-HCl, pH 7.6, 100 mM $MgCl_2$, 100 mM DTT, 500 ug BSA), and water to 10 ul final volume, and incubated for 4 hours at 16 °C. After ligation the cDNA is amplified using a linker specific primer. The PCR conditions are 35

A58

μl of water, 5 μl of Taq buffer (100 mM Tris-HCl, pH 8.3, 500 mM KCl, 15 mM $MgCl_2$, and 0.01% (w/v) gelatin), 1.5 μl 5 mM dNTP mix (equimolar mixture of dATP, dCTP, dGTP, dTTP with a concentration of 1.25 mM each dNTP), 2.5 μl of linker specific primers (10 pmol/μl), 2.5 μl of V_{HBACK} primers (10 pmol/μl), 2.5 μl of cDNA and overlay 2 drops of mineral oil. Heat to 94° C and add 1 U of Taq DNA polymerase. Amplify using 30 cycles of 94° C for 1 minute, 57° C for 1 minute, 72° C for 2 minutes. To the PCR reaction add 7.5M ammonium acetate to a final concentration of 2 M and precipitate the DNA by the addition of 1 volume of isopropanol and incubate at 25° C for 10 minutes. Pellet the DNA by centrifugation (13,000 rpm, 10 minutes) and dissolve the pellet in 100 μl of 0.3 M sodium acetate and reprecipitate by the addition of 2.5 volumes of ethanol. Incubate at -20° C for 30 minutes. Pellet the DNA by centrifugation (13,000 rpm, 10 minutes) and rinse the pellet with 70% ethanol. Dry the pellet *in vacuo* for 10 minutes then redissolve the dried pellets in 10-100 μl of TE buffer to 0.2-1.0 mg/ml. Determine the DNA concentration by absorbance at 260 nm.

A58 Cont.

Please replace the paragraph on page 108, lines 3-17, with the following paragraph:

The functional region "D" refer to sequences which are used to "divide" the library by providing sequences for specific PCR amplification. They are composed of a known sequences. An example is the sequence 5'-GATC(A)(T)GATC(G)TC(C)GA(A)G-3' SEQ ID No. 1 in which the positions in parenthesis vary. Oligonucleotides encoding the D sequences are designed to provide a minimum of sequence identity among each other and among known sequences in the database, to maximize specific amplification during the PCR. Incorporating these sequences in the tags enables the library to be divided by PCR amplification using primers that are specific for the various sequences. For example, if the library has been tagged with the above sequence, a primer containing the sequence 5'-GATC(A)(T)GATC(G)TC(C)GA(A)G-3' SEQ ID No. 2 specifically amplifies one group of tagged molecules; whereas a primer

A59

containing the sequence 5'-GATC(G)(G)GATC(A)TC(A)GA(A)G-3' SEQ ID No. 3 amplifies a different group of tagged molecules.

A59 CONT.

Please replace the paragraph beginning on page 109, line 18, through page 110, line 5, with the following paragraph:

EXAMPLE 4

Construction of recombinant antibody libraries

A. Preparation of recombinant antibodies

Recombinant antibody libraries are prepared by methods known to those of skill in the art (see, *e.g.*, Kay *et al.* (1996) *Phage Display of Peptides and Proteins: A Laboratory Manual*, Academic Press, San Diego); McCafferty *et al.* (1996) *Antibody engineering: A practical Approach*, Oxford University Press, Oxford). Functional antibody fragments can be created by genetic cloning and recombination of the variable heavy (V_H) chain and variable light (V_L) chain genes from a mouse or human. The V_H and V_L chain genes are cloned by reverse transcribing poly(A)RNA isolated from spleen tissue and then using specific primers to amplify the V_H and V_L chain genes by PCR. The V_H and V_L chain genes are joined by a linker region (a typical linker to produce a single-chain antibody fragment, scFv, includes DNA sequences encoding the amino acid sequence $(Gly_4Ser)_3$). After the V_H -linker-V_L genes have been assembled and amplified by PCR, the products are transcribed and translated directly or cloned into an expression plasmid and then expressed either *in vivo* or *in vitro*.

A60

Please replace the paragraph beginning on page 110, line 14, through page 111, line 2, with the following paragraph:

Amplification of the V_H and V_L chain genes is accomplished with sets of PCR primers that correspond to consensus sequences flanking these genes (McCafferty *et al.* (1996) *Antibody engineering: A practical Approach*, Oxford University Press, Oxford). In a 0.5 ml microcentrifuge tube mix the following: 35 µl of water, 5 µl of Taq buffer (100 mM Tris-HCl, pH 8.3, 500 mM KCl, 15 mM

A61

$MgCl_2$, and 0.01% (w/v) gelatin), 1.5 μl 5 mM dNTP mix (equimolar mixture of dATP, dCTP, dGTP, dTTP with a concentration of 1.25 mM each dNTP), 2.5 μl of FOR primers (10 pmol/μl), 2.5 μl of BACK primers (10 pmol/μl). The mixture is irradiated with UV light at 254 nm for 5 minutes. In a new 0.5 ml tube add 47.5 μl of the irradiated mix to 2.5 μl of cDNA and optionally overlay 2 drops of mineral oil. Heat to 94° C and add 1 U of Taq DNA polymerase. Amplify using 30 cycles of 94° C for 1 minute, 57° C for 1 minute, 72° C for 2 minutes. Isolate and purify the amplified DNA from the primers by electrophoresis in a low melting temperature agarose gel. Estimate the quantities of purified V_H and V_L chain DNA. For a mouse antibody library set up the following reaction: approximately 50 ng each of V_H and V_L chain DNA and linker DNA, 2.5 ul of Taq buffer, 2 μl of 5 mM dNTP mix, water up to 25 μl, and 1 U of Taq DNA polymerase (1U/μl). Amplify using 20 cycles of 94° C for 1.5 minute, 65° C for 3 minutes.

Please replace the paragraph on page 113, lines 1-14, with the following paragraph:

EXAMPLE 5

Creation and production of scFvs

The HFN7.1 hybridoma (HFN7.1 deposited under ATCC accession no. CRL-1606) and 10F7MN hybridomas (10F7MN deposited under ATCC accession no. HB-8162) are obtained from American Tissue type collection. The IgG produced by HFN7.1 recognizes human fibronectin, while the IgG produced by 10F7MN recognizes human glycophorin-MN. Cells are expanded by growth in culture (Covance, Richmond CA) and provided as a frozen pellet. Messenger RNA is prepared using the mRNA direct kit (Qiagen) according to the manufacturer's instructions. 500ng of purified mRNA is diluted to 25ng/μl in sterile RNAse free H_2O and denatured at 65°C for 10 minutes, then cooled on ice for 5 minutes. First strand cDNA is created using the reagents and methods described in the "Mouse scFv Module" (Amersham Pharmacia).

Please replace the paragraph beginning on page 115, line 18, through page 116, line 16, with the following paragraph:

For large scale preparation of purified scFv, osmotic shock fluid from an induced culture is reacted with a metal chelate to capture the polyhistidine tagged scFv. Briefly, a single colony representing the desired clone is inoculated into 400 ml of 2xYT containing 100μg/ml ampicillin and shaken at 250 rpm overnight at 37°C. The culture is diluted to 800 ml of 2xYT containing 0.1% arabinose and 100μg/ml ampicillin. This culture is now shaken at 250rpm for 4 hours at 30°C to allow expression of the scFv. Bacteria are pelleted at 3000x g at 4°C for 15 minutes, and resuspended in 20% sucrose, 20mM Tris-HCl, 2.5mM EDTA, pH 8.0 at 5.0 OD Units (absorbance at 600nm). Cells are incubated on ice for 20 minutes and then pelleted at 3000xg for 10 minutes at 4°C. The supernatant is removed and saved. Following resuspension in 20mM Tris-HCl, 2.5mM EDTA, pH 8.0 at 5.0 OD units, cells are incubated on ice for 10 minutes and then pelleted at 3000xg for 10 minutes at 4°C. The supernatant from this step is combined with the previous supernatant and NaCl, imidazole, and $MgCl_2$ are added to final concentrations of 1M, 10mM, and 10mM respectively. Nickel-nitriloacetic acid agarose beads (Ni-NTA, Qiagen) are stirred with the combined supernatants overnight at 4°C. The beads are collected with centrifugation at 3000xg for 10 minutes at 4°C, and resuspended in 50mM NaH_2PO_4, 20mM imidazole, 300mM NaCl, pH 8.0 and loaded into a column. After allowing the resin to pack and this wash buffer to flow through, the scFv is eluted with successive 0.5ml fractions of 50mM NaH_2PO_4, 250mM Imidazole, 300mM NaCl, 50mM EDTA, pH 8.0. Fractions are analyzed by SDS-PAGE and staining with GelCode Blue (Pierce-Endogen) and those containing sufficient quantities of scFv are pooled and dialyzed vs PBS overnight at 4°C. Purified scFv is quantified using a modified Lowry assay (Pierce-Endogen) according to the manufacturer's instructions and stored in PBS + 20% glycerol at -80°C until use.

A63

Please r place the paragraph on page 117, lines 1-8, with th following paragraph:

Anti-tag capture antibodies

For microarray analyses of scFv function and specificity, capture antibodies specific for hemagglutinin (HA.11, specific for the influenza virus hemagglutinin epitope YPYDVPDYA; Covance catalog # MMS-101P, lot # 139027002) and Myc (9E10, specific for the EQKLISEEDL amino acid region of the Myc oncoprotein; Covance catalog # MMS-150P, lot # 139048002) were used. A negative control mouse IgG antibody (FLOPC-21; Sigma catalog # M3645) was also included in these assays.

A67

Please replace the paragraphs beginning on page 120, line 11, through page 121, line 19, with the following paragraphs:

Basic protocol for antibody and antigen incubations: FAST slides and nitrocellulose filters printed with CytoSets™ capture antibodies

After printing CytoSets™ capture antibodies onto FAST slides or nitrocellulose filters, these support media were allowed to dry as described. Slides and filters were then blocked with BBSA-T, for 30 min to 1 h, at ambient temperature (filters) or 37°C (slides). All incubations were done on an orbital table (ambient temperature incubations) or in a shaking incubator (37°C incubations).

A68

Purified, recombinant cytokine antigen (contained in each kit) was then diluted to various concentrations (typically between 1-10 ng/ml) in BBSA-T. Slides or filters, containing CytoSets™ capture antibodies, were then incubated with this antigen solution at ambient temperature (filters) or 37°C (slides). Slides and filters were then washed three times with PBS-T, 3-5 min per wash, at ambient temperature. These slides and filters, containing capture antibody with bound antigen, were then incubated with detector antibody (contained in each kit) diluted 1:2500 in BBSA-T for 1 h, at ambient temperature (filters) or

37°C (slides). Slides and filters were then washed with PBS-T as described above.

These slides and filters, containing capture antibody, bound antigen, and bound detector antibody, were then incubated with streptavidin•HRPO (contained in each kit) diluted 1:2500 in BBSA-T for 1 h, at ambient temperature (filters) or 37°C (slides). Slides and filters were then washed with PBS-T as described above. The slides and filters were then developed and imaged as described below.

A68 cont.

Basic protocol for antibody and antigen incubations: FAST slides printed with anti-peptide tag capture antibodies

After printing anti-peptide tag capture antibodies onto FAST slides, the slides were allowed to dry as described. Slides were then blocked with BBSA-T, for 30 min to 1 h, at 37°C in a shaking incubator (37°C incubations).

Purified scFvs, containing peptide tags, were then diluted to various concentrations (typically between 0.1 and 100 ìg/ml) in BBSA-T. Slides containing anti-peptide tag capture antibodies were then incubated with this antigen solution for 1 h at 37°C. Slides were then washed three times with PBS-T, 3-5 min per wash, at ambient temperature.

Slides containing anti-peptide tag capture antibodies and bound scFvs were then incubated with biotinylated human fibronectin or biotinylated human glycophorin (as antigens) diluted to various concentrations (typically 1-10 ìg/ml) in BBSA-T, for 1 h at 37°C. Slides were then washed with PBS-T as described above.

IN THE ABSTRACT:

Please amend the abstract as follows (a marked-up copy of the amended abstract is attached to this Amendment):

ABSTRACT

A69

Provided herein are addressable collections of anti-tag capture agents, such as antibodies, that are used as tools for sorting proteins containing